ON THE LOCATION OF EIGENVALUES OF OFF-DIAGONAL CONSTANT MATRICES

by

Michel Gendreau

Cette recherche a été rendue possible grâce à une subvention du C.R.S.H. (#410-81-0722-RL), à une subvention du fonds F.C.A.C. (#83-AS-0026) ainsi qu'à une bourse de chercheur du Programme de soutien à l'emploi scientifique, Secrétariat à la Science et à la technologie du Gouvernement du Québec.

Ce rapport a été publié grâce à une subvention du fonds F.C.A.C. pour l'aide et le soutien à la recherche.

Université de Montréal

Centre de recherche sur les transports - Publication #350

Août 1984

Dépôt légal - 4e trimestre 1984

ABSTRACT

In this paper, we show that it is possible to derive tight bounds on the eigenvalues of real matrices whose off-diagonal entries are all equal. In most cases, these bounds are much sharper than those obtained by applying Gershgorin's Circle Theorem.

RÉSUMÉ

Dans cet article, nous montrons qu'il est possible de trouver de très bonnes bornes sur les valeurs propres des matrices réelles dont tous les éléments hors-diagonale sont égaux. Ces bornes représentent dans la plupart des cas une amélioration sensible sur les bornes que l'on obtient en appliquant le théorème de Gershgorin.

Off-diagonal constant matrices are square real matrices which can be written as the sum of a diagonal matrix and a constant matrix. Matrices of this type occur in the convergence analysis of algorithms used to solve the spatial equilibrium problem (see Zubieta [2]). To correctly assess the convergence properties of these algorithms, it is necessary to determine the value of some of the eigenvalues of a given off-diagonal constant matrix or at least tight bounds on these eigenvalues. When the diagonal elements of the matrix are widely spaced and relatively large compared to the off-diagonal elements, Gershgorin's Theorem and other elementary results of linear algebra yield in most cases satisfactory bounds on the eigenvalues of interest (see, for example, Noble and Daniel [1]). However, when the preceding conditions are not met, it is impossible to proceed with the convergence analysis. In this paper, we derive new bounds on the eigenvalues of off-diagonal constant matrices which are particularly sharp whenever Gershgorin's Theorem and related results are of little use.

Let A be an n x n off-diagonal matrix such that $a_{ii} = \alpha + d_i, i=1,\ldots,n$, and $a_{ij} = \alpha$, $i \neq j$, where α and d_i's are given scalars, and set $D \triangleq \{d_i, i=1,\ldots,n\}$. Let now E be the set of distinct values in D indexed in ascending order and m be the number of elements in E (so that we have $e_1 < e_2 < \ldots < e_m$). For each j in $\{1,\ldots,m\}$, let n_j be the number of occurences of e_j in D and I_j be the subset of indices i such that $d_i = e_j$ (note that $n_j = |I_j|, j=1,\ldots,m$). In the following, we suppose that $\alpha > 0$; when $\alpha < 0$, one can simply derive bounds on the eigenvalues of (-A) and use the fact that, if λ is an eigenvalue of (-A), then $(-\lambda)$ is an eigenvalue of A.

Let us first notice that the eigenvectors of A can be partitioned into two sets : those whose components sum to zero and those whose components do not sum to zero. We will now examine each of these sets separately.

Let $\underline{x}$ be an eigenvector of A such that $\sum_{i=1}^{n} x_i \neq 0$. Without loss of generality we may suppose that $\sum_{i=1}^{n} x_i = 1$. We know that for $\underline{x}$ to be an eigenvector, we must have $A\underline{x} = \lambda\underline{x}$ for some eigenvalue λ of A. It is easily seen that, due to the special structure of A, the vector equation $A\underline{x} = \lambda\underline{x}$ is equivalent to the system $d_i x_i + \alpha = \lambda x_i, i=1,\ldots,n$. This in turn is equivalent to

$$x_i = \alpha(\lambda - d_i)^{-1}, i=1,\ldots,n, \tag{1}$$

which implies

$$\sum_{i=1}^{n} (\lambda - d_i)^{-1} = \alpha^{-1}$$

$$\Leftrightarrow \sum_{j=1}^{m} n_j(\lambda - e_j)^{-1} = \alpha^{-1}. \tag{2}$$

It is fairly easy to see that the function $f(\lambda) \triangleq \sum_{j=1}^{m} n_j(\lambda - e_j)^{-1}$ is stricly monotone decreasing on any interval not containing any e_j, that $\lim_{\lambda\to-\infty} f(\lambda) = \lim_{\lambda\to+\infty} f(\lambda) = 0$, and that $\lim_{\lambda\to e_j^-} f(\lambda) = -\infty$ and $\lim_{\lambda\to e_j^+} f(\lambda) = +\infty$. From this, we can deduce that equation (2) possesses exactly m distinct roots $\lambda_j, j=1,\ldots,m$, and that $\lambda_j \in (e_j, e_{j+1}), j=1,\ldots,m-1$, and $\lambda_m \in (e_m, +\infty)$. To each of these roots, we can associate an eigenvector $\underline{x}^j$ whose components sum to 1 :

$$x_i^j = \alpha(\lambda_j - d_i)^{-1}, \quad i=1,\ldots,n. \tag{3}$$

Let us now consider the eigenvectors of A whose components sum to 0. For any e_j in E such that $n_j > 1$, it is easily verified that e_j is an eigenvalue of A and that the invariant subspace associated with e_j contains the set $X_j = \{\underline{x} \in \mathbb{R}^n \mid \sum_{i\in I_j} x_i = 0 \; ; \; x_i = 0, \; i \notin I_j\}$. From this, we can infer that the multiplicity of e_j is no less than $(n_j - 1)$. On the other hand,

the sum of the multiplicities of the e_j's can be no more than n - m, since we have already found that the m roots of equation (2) are eigenvalues of A.

Using the fact that $\sum_{j=1}^{m} (n_j - 1) = (\sum_{j=1}^{n} n_j) - m = n - m$, we are now in position to conclude that :

- each of the roots of equation (2) is an eigenvalue of multiplicity 1 ;
- each e_j such that $n_j > 1$ is an eigenvalue of multiplicity $(n_j - 1)$.

We have thus succeeded in locating fairly precisely all but one of the eigenvalues of A : the (m-1) first roots of equation (2) which belong to the intervals (e_j, e_{j+1}), $j=1,\ldots,m-1$, and the eigenvalues which are equal to some of the e_j's. We will now proceed to bound the largest eigenvalue of A, λ_m, which we already know to be in the interval $(e_m, +\infty)$. It is a well-known fact that the largest eigenvalue of any symmetric matrix is larger than or equal to the largest diagonal element of this matrix ; therefore $\lambda_m > \max_i(d_i + \alpha) = e_m + \alpha$. As for an upper bound on λ_m, we can derive a crude one by using Gershgorin's Theorem : $\lambda_m \leq e_m + n\alpha$. This bound can be improved by defining a simple fixed-point scheme. From (3), we know that $x_i^m = \alpha(\lambda_m - d_i)^{-1}$, $i=1,\ldots,n$. Going back to the system $d_i x_i^m + \alpha = \lambda_m x_i^m$, $i=1,\ldots,m$, we may write $\lambda_m x_i^m = e_m x_i^m + \alpha + (d_i - e_m) x_i^m$, $i=1,\ldots,n$. Summing up these n equations, we get

$$\lambda_m (\sum_{i=1}^{n} x_i^m) = e_m (\sum_{i=1}^{n} x_i^m) + n\alpha + \sum_{i=1}^{n} (d_i - e_m) x_i^m,$$

which can be rewritten as

$$\lambda_m = e_m + n\alpha - \sum_{i=1}^{n} (e_m - d_i) x_i^m,$$

since $\sum_{i=1}^{n} x_i^m = 1$, or as

$$\lambda_m = e_m + n\alpha - \alpha \sum_{i=1}^{n} (e_m - d_i) (\lambda_m - d_i)^{-1}.$$

Using any known upper bound on λ_m as B^0 (for instance, the one obtained earlier), the iteration

$$B^{k+1} = e_m + n\alpha - \alpha \sum_{i=1}^{n} (e_m - d_i)\ (B^k - d_i)^{-1}, k=0,1,\ldots,$$

will produce a sequence $\{B^k\}_{k=0}^{\infty}$ of increasingly good upper bounds on λ_m.

Linear convergence of the sequence $\{B^k\}_{k=0}^{\infty}$ to λ_m is easily shown as follows.

For $B > e_m$, let $g(B) = e_m + n\alpha - \alpha \sum_{i=1}^{n} (e_m - d_i)\ (B - d_i)^{-1}$;

then $g'(B) = \alpha \sum_{i=1}^{n} (e_m - d_i)\ (B - d_i)^{-2}$.

When $B > \lambda_m > e_m$,

$$(e_m - d_i)\ (B - d_i)^{-1}\ \frac{e_m - e_1}{B - e_1} < \frac{e_m - e_1}{\lambda_m - e_1}\ , i=1,\ldots,n,$$

and $$\alpha\ (B - d_i)^{-1} < \alpha\ (\lambda_m - d_i)^{-1} = x_i^m\ , i=1,\ldots,n,$$

therefore

$$g'(B) < \frac{e_m - e_1}{\lambda_m - e_1} - \sum_{i=1}^{n} x_i^m\ \frac{e_m - e_1}{\lambda_m - e_1}\ , \text{ since } \sum_{i=1}^{n} x_i^m = 1.$$

By using the fact that $g(\lambda_m) = \lambda_m$, we can now write

$$g(B^k) - g(\lambda_m) = \int_{\lambda_m}^{B^k} g'(B)dB < \frac{e_m - e_1}{\lambda_m - e_1}\ (B^k - \lambda_m)$$

$$\Leftrightarrow \quad B^{k+1} - \lambda_m < \frac{e_m - e_1}{\lambda_m - e_1} \; (B^k - \lambda_m),$$

thus proving linear convergence of $\{B^k\}_{k=0}^{\infty}$ to λ_m with convergence ratio $\dfrac{e_m - e_1}{\lambda_m - e_1}$.

Our experience with small examples seems to indicate that, if one uses the crude bound found by applying Gershgorin's Theorem as B^0, only one iteration of the proposed fixed-point scheme might be necessary in order to get a satisfactory bound on λ_m.

ACKNOWLEDGEMENTS

I would like to thank Patrice Marcotte for his comments on an earlier version of this paper.

BIBLIOGRAPHY

[1] NOBLE, B. and DANIEL, J.W., Applied Linear Algebra. Prentice-Hall, Englewood Cliffs, N.J., 1969.

[2] ZUBIETA, L. Algorithmes pour le modèle d'équilibre spatial. Ph.D. thesis. Département d'informatique et de recherche opérationnelle, Université de Montréal, Forthcoming.